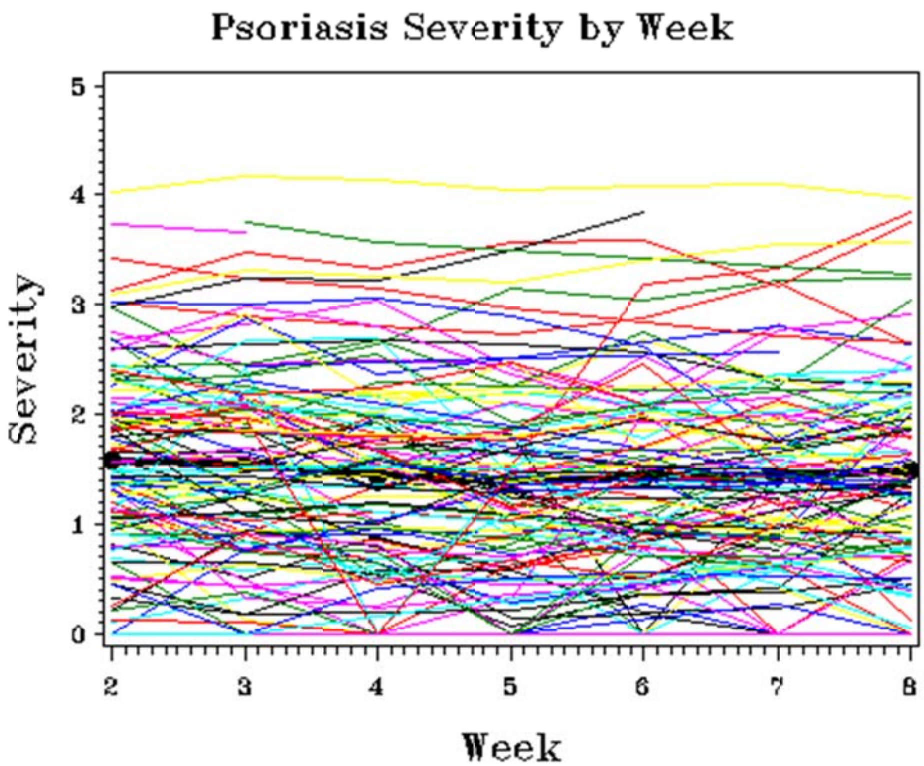
Within-Person Fluctuation in Symptom Severity over Time



These data come from a study of weekly fluctuation in psoriasis severity. There was no intervention and no real reason to expect systematic growth (as shown by the mostly flat line for the means).

But we still need to be concerned about the model for the variance. Specifically, the variances across occasions may need to differ, and the covariances and correlations across occasions may need to differ as well.

We will begin by including just a fixed intercept in the model for the means, and examine different models for the variance instead.

```
/* PROC MIXED syntax commands used below:

PROC MIXED:
      DATA gives file used for analysis
      METHOD=REML specifies restricted ML (REML) or ML (REML is default)
      COVTEST gives significance tests of covariance parameters
      NOCLPRINT suppresses class variable list
      NOITPRINT suppresses iteration history (but use for troubleshooting)
      MAXITER=1000 increases #iterations to 1000

CLASS: defines categorical variables
       Observation identifiers (subid, week) are categorical

MODEL: defines equation, where [outcome] = [predictors]
       / SOLUTION asks for parameter estimates
       DDFM=Satterthwaite specifies Satterthwaite denominator DF

RANDOM: lists random effects (INTERCEPT is not included by default)
        TYPE=UN specifies unstructured matrix (all variances & covariances)
        / SUBJECT=subid defines nesting structure for level 2 persons
        / G, GCORR prints covariance, correlation matrix of random effects
        / V, VCORR prints combined covariance, correlation matrix from G and R

REPEATED: allows structuring of the residual variance and repeated measures
          week indicates level-1 nesting (weeks within persons)
          / TYPE= specifies covariance structure (see SAS manual for list)
          / R, RCORR asks for covariance and correlation matrices
          REPEATED is included by default for random effects models
          with TYPE=VC (is diagonal matrix --> equal variances no covariances)
*/
```

```
* Defining data file as macro variable to be replaced throughout;
%LET datafile=p2stack;
```

Let's begin by testing the most basic model for the variance: an e-only model, which assumes equal variance over persons and occasions and that all observations are independent—it is a cross-sectional model.

```
TITLE 'SAS E-only R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / TYPE=VC R RCORR SUBJECT=subid;
RUN;
```

```
            Dimensions
Covariance Parameters              1    Cov Parms = total number of parameters in model for variances
Columns in X                       1    Columns in X = total number of fixed effects plus intercept
Columns in Z                       0    Columns in Z = total number of U's (not counting covariances)
Subjects                         124    Subjects = number of persons in level 2 (with at least 1 obs)
Max Obs Per Subject                7    Max Obs Per Subject = max number of time points per person

          Number of Observations
Number of Observations Read           770
Number of Observations Used           770
Number of Observations Not Used         0    Make sure to keep track of how many cases get dropped!!

                     Estimated R Matrix for SUBID 100 = TOTAL variance
 Row       Col1        Col2        Col3        Col4        Col5        Col6        Col7
   1     0.7820
   2                 0.7820
   3                             0.7820
   4                                         0.7820
   5                                                     0.7820
   6                                                                 0.7820
   7                                                                             0.7820

                    Estimated R Correlation Matrix for SUBID 100
 Row       Col1        Col2        Col3        Col4        Col5        Col6        Col7
   1     1.0000
   2                 1.0000
   3                             1.0000
   4                                         1.0000
   5                                                     1.0000
   6                                                                 1.0000
   7                                                                             1.0000

                Covariance Parameter Estimates
Cov                             Standard         Z
Parm     Subject    Estimate       Error     Value        Pr Z

WEEK     SUBID        0.7820     0.03988     19.61      <.0001  the total variance (equal over occasions)

           Fit Statistics
-2 Res Log Likelihood          1999.9
AIC (smaller is better)        2001.9
AICC (smaller is better)       2001.9
BIC (smaller is better)        2004.7

                  Solution for Fixed Effects
                             Standard
Effect        Estimate         Error      DF    t Value     Pr > |t|
Intercept       1.4583       0.03187     769      45.76       <.0001  fixed intercept → here, the grand mean
```

Next, we can test the **compound symmetry** model, which assumes equal total variance over time and equal total covariance over time.

```
TITLE 'SAS Compound Symmetry R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / TYPE=CS R RCORR SUBJECT=subid;
RUN;
```

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z	0
Subjects	124
Max Obs Per Subject	7

Estimated R Matrix for SUBID 100 → TOTAL variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8127	0.6820	0.6820	0.6820	0.6820	0.6820	0.6820
2	0.6820	0.8127	0.6820	0.6820	0.6820	0.6820	0.6820
3	0.6820	0.6820	0.8127	0.6820	0.6820	0.6820	0.6820
4	0.6820	0.6820	0.6820	0.8127	0.6820	0.6820	0.6820
5	0.6820	0.6820	0.6820	0.6820	0.8127	0.6820	0.6820
6	0.6820	0.6820	0.6820	0.6820	0.6820	0.8127	0.6820
7	0.6820	0.6820	0.6820	0.6820	0.6820	0.6820	0.8127

Estimated R Correlation Matrix for SUBID 100 → TOTAL correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393
2	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393	0.8393
3	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393	0.8393
4	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393	0.8393
5	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393	0.8393
6	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000	0.8393
7	0.8393	0.8393	0.8393	0.8393	0.8393	0.8393	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
CS	SUBID	0.6820	0.09019	7.56	<.0001	compound symmetry parameter
Residual		0.1306	0.007270	17.97	<.0001	residual variance (Var of e's)

Fit Statistics

-2 Res Log Likelihood	1049.7
AIC (smaller is better)	1053.7
AICC (smaller is better)	1053.7
BIC (smaller is better)	1059.4

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	950.14	<.0001

Is this CS model better than the e-only model? How do we know?

$-2\Delta LL(1) = 950.2$, $p < 001$ (critical value is 2.71 because random intercept variance/CS parameter is bounded at 0)

Note that the ICC is given in the RCORR matrix (.8393).

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.4767	0.07548	123	19.56	<.0001

Note the large change in the SE and df for testing the fixed intercept…

Next, let's see what the observed pattern of variances and covariances over time looks like by estimating an *n*-order unstructured **R** baseline model for the variances → each variance and covariance estimated is separately, with no constraints, so this is a description, not really a model. Note: you can only estimate an unstructured model if time is balanced across persons and you have more people than parameters estimated!

```
TITLE 'SAS Unstructured n-Order R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / R RCORR TYPE=UN SUBJECT=subid; RUN;
```

Dimensions

Covariance Parameters	28
Columns in X	1
Columns in Z	0
Subjects	124
Max Obs Per Subject	7

Note that given n=7, this model requires $n*(n+1) / 2 = 28$ covariance parameters!! You'd need at least 28 people to estimate it (and it may not be possible even then).

Estimated R Matrix for SUBID 100 → TOTAL variance and covariance → ANSWER KEY

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7477**	0.7190	0.6817	0.6807	0.6124	0.6182	0.6328
2	0.7190	**0.8266**	0.7428	0.7475	0.6571	0.6619	0.6541
3	0.6817	0.7428	**0.8346**	0.7498	0.6524	0.6390	0.6172
4	0.6807	0.7475	0.7498	**0.8366**	0.7292	0.7206	0.6805
5	0.6124	0.6571	0.6524	0.7292	**0.8285**	0.7280	0.7313
6	0.6182	0.6619	0.6390	0.7206	0.7280	**0.7911**	0.7352
7	0.6328	0.6541	0.6172	0.6805	0.7313	0.7352	**0.8464**

Estimated R Correlation Matrix for SUBID 100 → TOTAL correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.9146	0.8630	0.8608	0.7781	0.8038	0.7955
2	**0.9146**	1.0000	0.8943	0.8989	0.7940	0.8185	0.7820
3	**0.8630**	**0.8943**	1.0000	0.8974	0.7846	0.7864	0.7343
4	**0.8608**	**0.8989**	**0.8974**	1.0000	0.8760	0.8859	0.8088
5	**0.7781**	**0.7940**	**0.7846**	**0.8760**	1.0000	0.8993	0.8734
6	**0.8038**	**0.8185**	**0.7864**	**0.8859**	**0.8993**	1.0000	0.8985
7	**0.7955**	**0.7820**	**0.7343**	**0.8088**	**0.8734**	**0.8985**	1.0000

Covariance Parameter Estimates table is too big to print (has 28 entries!)...

Fit Statistics

-2 Res Log Likelihood	893.1
AIC (smaller is better)	949.1
AICC (smaller is better)	951.3
BIC (smaller is better)	1028.1

The fit statistics for this n-order unstructured **R** (and completely unparsimonious model) will serve as a baseline with which to compare more parsimonious models for the variances.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5178	0.07429	122	20.43	<.0001

Ok, so this is what we are trying to reproduce… next we will try a few **R**-only models and see how they fit.

AR(1) **R**-Only Model

```
TITLE 'SAS AR1 R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / TYPE=AR(1) R RCORR SUBJECT=subid; RUN;
```

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z	0
Subjects	124
Max Obs Per Subject	7

Estimated R Matrix for SUBID 100 → TOTAL variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7999**	0.7163	0.6415	0.5745	0.5145	0.4607	0.4126
2	0.7163	**0.7999**	0.7163	0.6415	0.5745	0.5145	0.4607
3	0.6415	0.7163	**0.7999**	0.7163	0.6415	0.5745	0.5145
4	0.5745	0.6415	0.7163	**0.7999**	0.7163	0.6415	0.5745
5	0.5145	0.5745	0.6415	0.7163	**0.7999**	0.7163	0.6415
6	0.4607	0.5145	0.5745	0.6415	0.7163	**0.7999**	0.7163
7	0.4126	0.4607	0.5145	0.5745	0.6415	0.7163	**0.7999**

Estimated R Correlation Matrix for SUBID 100 → TOTAL correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8955	0.8020	0.7182	0.6432	0.5760	0.5158
2	**0.8955**	1.0000	0.8955	0.8020	0.7182	0.6432	0.5760
3	**0.8020**	**0.8955**	1.0000	0.8955	0.8020	0.7182	0.6432
4	**0.7182**	**0.8020**	**0.8955**	1.0000	0.8955	0.8020	0.7182
5	**0.6432**	**0.7182**	**0.8020**	**0.8955**	1.0000	0.8955	0.8020
6	**0.5760**	**0.6432**	**0.7182**	**0.8020**	**0.8955**	1.0000	0.8955
7	**0.5158**	**0.5760**	**0.6432**	**0.7182**	**0.8020**	**0.8955**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
AR(1)	SUBID	0.8955	0.01148	78.01	<.0001	Auto-regressive parameter
Residual		0.7999	0.08044	9.94	<.0001	Total variance (equal across weeks)

Fit Statistics

-2 Res Log Likelihood	990.9
AIC (smaller is better)	994.9
AICC (smaller is better)	994.9
BIC (smaller is better)	1000.5

AR1 also forces all variances equal, but with correlations declining sharply with time (r, r^2, r^3... = .8955, .8020, .7182…). The BIC (but not the AIC) is happier with this constraint.

The AR1 model fits significantly worse than the UN model, $-2\Delta LL(26) = 97.8$, $p < .001$.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5148	0.07031	129	21.54	<.0001

Toeplitz(*n*) **R**-Only Model

```
TITLE 'SAS Toeplitz (n=7 Bands) R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / TYPE=TOEP(7) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → TOTAL variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.8103**	0.7257	0.6974	0.6576	0.6394	0.6368	0.6541
2	0.7257	**0.8103**	0.7257	0.6974	0.6576	0.6394	0.6368
3	0.6974	0.7257	**0.8103**	0.7257	0.6974	0.6576	0.6394
4	0.6576	0.6974	0.7257	**0.8103**	0.7257	0.6974	0.6576
5	0.6394	0.6576	0.6974	0.7257	**0.8103**	0.7257	0.6974
6	0.6368	0.6394	0.6576	0.6974	0.7257	**0.8103**	0.7257
7	0.6541	0.6368	0.6394	0.6576	0.6974	0.7257	**0.8103**

Estimated R Correlation Matrix for SUBID 100 → TOTAL correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859	0.8072
2	**0.8955**	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859
3	**0.8606**	**0.8955**	1.0000	0.8955	0.8606	0.8115	0.7890
4	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955	0.8606	0.8115
5	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955	0.8606
6	**0.7859**	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000	0.8955
7	**0.8072**	**0.7859**	**0.7890**	**0.8115**	**0.8606**	**0.8955**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
TOEP(2)	SUBID	0.7257	0.09004	8.06	<.0001	Covariance for t-1
TOEP(3)	SUBID	0.6974	0.08997	7.75	<.0001	Covariance for t-2
TOEP(4)	SUBID	0.6576	0.09003	7.30	<.0001	Covariance for t-3
TOEP(5)	SUBID	0.6394	0.09019	7.09	<.0001	Covariance for t-4
TOEP(6)	SUBID	0.6368	0.09066	7.02	<.0001	Covariance for t-5
TOEP(7)	SUBID	0.6541	0.09179	7.13	<.0001	Covariance for t-6
Residual		0.8103	0.09014	8.99	<.0001	Total variance (equal across weeks)

Fit Statistics

-2 Res Log Likelihood	926.0
AIC (smaller is better)	940.0
AICC (smaller is better)	940.1
BIC (smaller is better)	959.7

The **Toeplitz(*n*)** model estimates a separate covariance for each time lag. The AIC and BIC are the lowest yet. Also, the TOEP(*n*) model fits almost not significantly worse than the n-order UN model, $-2\Delta LL(21) = 32.9$, $p = .047$.

Could this be our model? Does it need to have heterogeneous variances (separate variances per week) to fit better?

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.4884	0.07533	122	19.76	<.0001

Heterogeneous Variance Toeplitz(*n*) R-Only Model

```
TITLE 'SAS Heterogeneous Toeplitz (n=7 bands) R Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      REPEATED week / TYPE=TOEPH(7) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → TOTAL variance and covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7346**	0.6916	0.6896	0.6409	0.6363	0.5985	0.6332
2	0.6916	**0.8077**	0.7512	0.7126	0.6866	0.6322	0.6485
3	0.6896	0.7512	**0.8665**	0.7667	0.7540	0.6738	0.6766
4	0.6409	0.7126	0.7667	**0.8414**	0.7718	0.7039	0.6861
5	0.6363	0.6866	0.7540	0.7718	**0.8781**	0.7470	0.7431
6	0.5985	0.6322	0.6738	0.7039	0.7470	**0.7882**	0.7314
7	0.6332	0.6485	0.6766	0.6861	0.7431	0.7314	**0.8416**

Note that the *covariances are unequal* within a band (because the variances are allowed to differ over time), but the *correlations are equal* within a band.

Estimated R Correlation Matrix for SUBID 100 → TOTAL correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8979	0.8644	0.8153	0.7923	0.7865	0.8053
2	**0.8979**	1.0000	0.8979	0.8644	0.8153	0.7923	0.7865
3	0.8644	**0.8979**	1.0000	0.8979	0.8644	0.8153	0.7923
4	0.8153	0.8644	**0.8979**	1.0000	0.8979	0.8644	0.8153
5	0.7923	0.8153	0.8644	**0.8979**	1.0000	0.8979	0.8644
6	0.7865	0.7923	0.8153	0.8644	**0.8979**	1.0000	0.8979
7	0.8053	0.7865	0.7923	0.8153	0.8644	**0.8979**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
Var(1)	SUBID	0.7346	0.09397	7.82	<.0001	total variance at week1
Var(2)	SUBID	0.8077	0.1016	7.95	<.0001	total variance at week2
Var(3)	SUBID	0.8665	0.1129	7.68	<.0001	total variance at week3
Var(4)	SUBID	0.8414	0.1115	7.55	<.0001	total variance at week4
Var(5)	SUBID	0.8781	0.1162	7.56	<.0001	total variance at week5
Var(6)	SUBID	0.7882	0.1030	7.65	<.0001	total variance at week6
Var(7)	SUBID	0.8416	0.1092	7.70	<.0001	total variance at week7
TOEPH(1)	SUBID	0.8979	0.01292	69.51	<.0001	correlation for t-1
TOEPH(2)	SUBID	0.8644	0.01744	49.57	<.0001	correlation for t-2
TOEPH(3)	SUBID	0.8153	0.02429	33.56	<.0001	correlation for t-3
TOEPH(4)	SUBID	0.7923	0.02790	28.40	<.0001	correlation for t-4
TOEPH(5)	SUBID	0.7865	0.03056	25.73	<.0001	correlation for t-5
TOEPH(6)	SUBID	0.8053	0.03340	24.11	<.0001	correlation for t-6

Fit Statistics

-2 Res Log Likelihood	921.5
AIC (smaller is better)	947.5
AICC (smaller is better)	947.9
BIC (smaller is better)	984.1

The homogeneous variance Toeplitz model is nested within the heterogeneous variance Toeplitz model, so we can compare deviances: $-2\Delta LL(6) = 4.5$, $p = .609$. Nope, separate variances doesn't help the model fit better (AIC and BIC agree).

The TOEPH(n) model fits significantly worse than the UN(n) model, $-2\Delta LL(15) = 28.4$, $p = .019$.

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5204	0.07417	124	20.50	<.0001

Another possibility: Just as we'll see in models for change, models for within-person fluctuation can combine the **G** matrix (for between-person random effects) and the **R** matrix (for within-person residuals) to recapture the total variance-covariance matrix. In this case, random slopes for time won't be helpful, but a random intercept might be. Adding a random intercept variance would allow the separation of between-person and within-person variance and remove the primary source of covariance from the **R** matrix. Thus, we will put the random intercept variance in **G**, and then try to find the pattern of what is left in **R**.

First, let's see the structure of just the RESIDUAL variances (after removing random intercept):

Random Intercept in **G** with UN(n−1) in **R** Model → Have to eliminate last covariance for the model to be identified because there is only one lag6 covariance (T1 with T7), so it is not separately estimable.

```
TITLE "SAS Random Intercept + Unstructured(n-1) R Model";
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / TYPE=UN V VCORR G SUBJECT=subid;
      REPEATED week / TYPE=UN(6) R RCORR SUBJECT=subid; RUN;
```

Dimensions

Covariance Parameters	29	Note that it still counts the one we left out...
Columns in X	1	
Columns in Z Per Subject	1	
Subjects	124	
Max Obs Per Subject	7	

Estimated R Matrix for SUBID 100 → WP residual variances and covariances

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1148**	0.08620	0.04892	0.04793	-0.02044	-0.01461	
2	0.08620	**0.1938**	0.1100	0.1147	0.02429	0.02907	0.02126
3	0.04892	0.1100	**0.2018**	0.1170	0.01962	0.006156	-0.01561
4	0.04793	0.1147	0.1170	**0.2038**	0.09643	0.08783	0.04774
5	-0.02044	0.02429	0.01962	0.09643	**0.1957**	0.09523	0.09853
6	-0.01461	0.02907	0.006156	0.08783	0.09523	**0.1583**	0.1024
7		0.02126	-0.01561	0.04774	0.09853	0.1024	**0.2136**

Estimated R Correlation Matrix for SUBID 100 → WP residual correlations

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5778	0.3213	0.3133	-0.1364	-0.1084	
2	0.5778	1.0000	0.5563	0.5773	0.1247	0.1660	0.1045
3	0.3213	0.5563	1.0000	0.5771	0.09875	0.03445	-0.07519
4	0.3133	0.5773	0.5771	1.0000	0.4830	0.4891	0.2288
5	-0.1364	0.1247	0.09875	0.4830	1.0000	0.5412	0.4820
6	-0.1084	0.1660	0.03445	0.4891	0.5412	1.0000	0.5571
7		0.1045	-0.07519	0.2288	0.4820	0.5571	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6328

Estimated V Matrix for SUBID 100 → TOTAL variance and covariance (matches previous Unstructured R only)

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.7477	0.7190	0.6817	0.6807	0.6124	0.6182	0.6328
2	0.7190	0.8266	0.7428	0.7475	0.6571	0.6619	0.6541
3	0.6817	0.7428	0.8346	0.7498	0.6524	0.6390	0.6172
4	0.6807	0.7475	0.7498	0.8366	0.7292	0.7206	0.6805
5	0.6124	0.6571	0.6524	0.7292	0.8285	0.7280	0.7313
6	0.6182	0.6619	0.6390	0.7206	0.7280	0.7911	0.7352
7	0.6328	0.6541	0.6172	0.6805	0.7313	0.7352	0.8464

Fit Statistics

-2 Res Log Likelihood	893.1
AIC (smaller is better)	949.1
AICC (smaller is better)	951.3
BIC (smaller is better)	1028.1

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.5178	0.07429	122	20.43	<.0001

We'll try two types of residual correlations: AR1 and Toeplitz (lag 5, 4, 3, 2, and 1).

Random Intercept in **G** with AR(1) in **R** Model

```
TITLE 'SAS Random Intercept + AR1 Correlation Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / V VCORR G TYPE=UN SUBJECT=subid;
      REPEATED week / TYPE=AR(1) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → WP residual variance & covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1731**	0.08887	0.04564	0.02343	0.01203	0.006180	0.003173
2	0.08887	**0.1731**	0.08887	0.04564	0.02343	0.01203	0.006180
3	0.04564	0.08887	**0.1731**	0.08887	0.04564	0.02343	0.01203
4	0.02343	0.04564	0.08887	**0.1731**	0.08887	0.04564	0.02343
5	0.01203	0.02343	0.04564	0.08887	**0.1731**	0.08887	0.04564
6	0.006180	0.01203	0.02343	0.04564	0.08887	**0.1731**	0.08887
7	0.003173	0.006180	0.01203	0.02343	0.04564	0.08887	**0.1731**

Estimated R Correlation Matrix for SUBID 100 → WP residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5135	0.2637	0.1354	0.06954	0.03571	0.01834
2	**0.5135**	1.0000	0.5135	0.2637	0.1354	0.06954	0.03571
3	**0.2637**	**0.5135**	1.0000	0.5135	0.2637	0.1354	0.06954
4	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135	0.2637	0.1354
5	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135	0.2637
6	**0.03571**	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000	0.5135
7	**0.01834**	**0.03571**	**0.06954**	**0.1354**	**0.2637**	**0.5135**	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6350

Estimated V Matrix for SUBID 100 → TOTAL variance and covariance put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.8080**	0.7238	0.6806	0.6584	0.6470	0.6411	0.6381
2	0.7238	**0.8080**	0.7238	0.6806	0.6584	0.6470	0.6411
3	0.6806	0.7238	**0.8080**	0.7238	0.6806	0.6584	0.6470
4	0.6584	0.6806	0.7238	**0.8080**	0.7238	0.6806	0.6584
5	0.6470	0.6584	0.6806	0.7238	**0.8080**	0.7238	0.6806
6	0.6411	0.6470	0.6584	0.6806	0.7238	**0.8080**	0.7238
7	0.6381	0.6411	0.6470	0.6584	0.6806	0.7238	**0.8080**

Estimated V Correlation Matrix for SUBID 100 → TOTAL correlation put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8958	0.8423	0.8148	0.8007	0.7935	0.7898
2	**0.8958**	1.0000	0.8958	0.8423	0.8148	0.8007	0.7935
3	**0.8423**	**0.8958**	1.0000	0.8958	0.8423	0.8148	0.8007
4	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958	0.8423	0.8148
5	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958	0.8423
6	**0.7935**	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000	0.8958
7	**0.7898**	**0.7935**	**0.8007**	**0.8148**	**0.8423**	**0.8958**	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6350	0.09026	7.04	<.0001	Random Intercept variance (Var of U_0's)
AR(1)	SUBID	0.5135	0.05179	9.92	<.0001	The auto-regressive correlation in R
Residual		0.1731	0.01777	9.74	<.0001	Residual e variance (equal across weeks)

Fit Statistics	
-2 Res Log Likelihood	943.8
AIC (smaller is better)	949.8
AICC (smaller is better)	949.9
BIC (smaller is better)	958.3

> The AIC doesn't like this one better than our best model so far (homogeneous Toeplitz), but the BIC could go either way. Although adding the random intercept variance is a significant improvement over the R-only AR1, but this still fits worse than UN(*n*-1).

Random Intercept in **G** with Toeplitz(6) in **R** Model (allows 5 lagged covariances; sets the last one to 0)

```
TITLE 'SAS Random Intercept + Lag-5 Covariance Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / V VCORR G TYPE=UN SUBJECT=subid;
      REPEATED week / TYPE=TOEP(6) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → WP residual variance & covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1562**	0.07152	0.04321	0.003458	-0.01479	-0.01730	
2	0.07152	**0.1562**	0.07152	0.04321	0.003458	-0.01479	-0.01730
3	0.04321	0.07152	**0.1562**	0.07152	0.04321	0.003458	-0.01479
4	0.003458	0.04321	0.07152	**0.1562**	0.07152	0.04321	0.003458
5	-0.01479	0.003458	0.04321	0.07152	**0.1562**	0.07152	0.04321
6	-0.01730	-0.01479	0.003458	0.04321	0.07152	**0.1562**	0.07152
7		-0.01730	-0.01479	0.003458	0.04321	0.07152	**0.1562**

Estimated R Correlation Matrix for SUBID 100 → WP residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.4579	0.2766	0.02214	-0.09467	-0.1108	
2	0.4579	1.0000	0.4579	0.2766	0.02214	-0.09467	-0.1108
3	0.2766	0.4579	1.0000	0.4579	0.2766	0.02214	-0.09467
4	0.02214	0.2766	0.4579	1.0000	0.4579	0.2766	0.02214
5	-0.09467	0.02214	0.2766	0.4579	1.0000	0.4579	0.2766
6	-0.1108	-0.09467	0.02214	0.2766	0.4579	1.0000	0.4579
7		-0.1108	-0.09467	0.02214	0.2766	0.4579	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6541

Estimated V Matrix for SUBID 100 → TOTAL variance and covariance put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	0.8104	0.7257	0.6974	0.6576	0.6394	0.6368	0.6541
2	0.7257	0.8104	0.7257	0.6974	0.6576	0.6394	0.6368
3	0.6974	0.7257	0.8104	0.7257	0.6974	0.6576	0.6394
4	0.6576	0.6974	0.7257	0.8104	0.7257	0.6974	0.6576
5	0.6394	0.6576	0.6974	0.7257	0.8104	0.7257	0.6974
6	0.6368	0.6394	0.6576	0.6974	0.7257	0.8104	0.7257
7	0.6541	0.6368	0.6394	0.6576	0.6974	0.7257	0.8104

Estimated V Matrix for SUBID 100 → TOTAL correlation put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859	0.8072
2	0.8955	1.0000	0.8955	0.8606	0.8115	0.7890	0.7859
3	0.8606	0.8955	1.0000	0.8955	0.8606	0.8115	0.7890
4	0.8115	0.8606	0.8955	1.0000	0.8955	0.8606	0.8115
5	0.7890	0.8115	0.8606	0.8955	1.0000	0.8955	0.8606
6	0.7859	0.7890	0.8115	0.8606	0.8955	1.0000	0.8955
7	0.8072	0.7859	0.7890	0.8115	0.8606	0.8955	1.0000

```
                    Covariance Parameter Estimates
                                       Standard         Z
Cov Parm      Subject     Estimate        Error     Value        Pr Z
UN(1,1)       SUBID         0.6541      0.09179      7.13      <.0001  Random Intercept U0 Variance
TOEP(2)       SUBID        0.07152      0.02189      3.27      0.0011  Residual covariance for t-1
TOEP(3)       SUBID        0.04321      0.02130      2.03      0.0425  Residual covariance for t-2
TOEP(4)       SUBID       0.003458      0.02096      0.16      0.8690  Residual covariance for t-3
TOEP(5)       SUBID       -0.01479      0.01860     -0.80      0.4266  Residual covariance for t-4
TOEP(6)       SUBID       -0.01730      0.01487     -1.16      0.2444  Residual covariance for t-5
Residual                    0.1562      0.02178      7.17      <.0001  Residual e Variance (equal across weeks)

              Fit Statistics

-2 Res Log Likelihood             926.0
AIC (smaller is better)           940.0
AICC (smaller is better)          940.1
BIC (smaller is better)           959.7
```

This RI + TOEP(n-1) 5-lag model is equivalent to the Toeplitz(n) 6-lag model without the random intercept, as shown by the fit statistics.

But this RI+5lag model has an interpretational advantage: rather than testing whether the total lagged covariance (in **V**) is different than 0, it tests whether the lagged covariance (in **R**) is different from 0 *after taking out individual differences in level*. It looks like only some of the lags are significant, so we can probably simplify the model without hurting fit.

Summary of sequential models taking out lagged residual covariances:

TOEP(5) → 4 lags only: REML −2LL = 927.2 AIC = 939.2 BIC = 956.2
TOEP(4) → 3 lags only: REML −2LL = 927.3 AIC = 937.3 BIC = 951.4 ← Best so far
TOEP(3) → 2 lags only: REML −2LL = 933.6 AIC = 941.6 BIC = 952.8

```
TITLE 'SAS Random Intercept + 3-Lag TOEP Covariance Model - THE WINNER SO FAR';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / V VCORR G TYPE=UN SUBJECT=subid;
      REPEATED week / TYPE=TOEP(4) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → WP residual variance & covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1702**	0.08576	0.05791	0.01872			
2	0.08576	**0.1702**	0.08576	0.05791	0.01872		
3	0.05791	0.08576	**0.1702**	0.08576	0.05791	0.01872	
4	0.01872	0.05791	0.08576	**0.1702**	0.08576	0.05791	0.01872
5		0.01872	0.05791	0.08576	**0.1702**	0.08576	0.05791
6			0.01872	0.05791	0.08576	**0.1702**	0.08576
7				0.01872	0.05791	0.08576	**0.1702**

Estimated R Correlation Matrix for SUBID 100 → WP residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5038	0.3402	0.1100			
2	0.5038	1.0000	0.5038	0.3402	0.1100		
3	0.3402	0.5038	1.0000	0.5038	0.3402	0.1100	
4	0.1100	0.3402	0.5038	1.0000	0.5038	0.3402	0.1100
5		0.1100	0.3402	0.5038	1.0000	0.5038	0.3402
6			0.1100	0.3402	0.5038	1.0000	0.5038
7				0.1100	0.3402	0.5038	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6395

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6395	0.09004	7.10	<.0001	Random Intercept U Variance
TOEP(2)	SUBID	0.08576	0.01267	6.77	<.0001	Residual covariance for t-1
TOEP(3)	SUBID	0.05791	0.009852	5.88	<.0001	Residual covariance for t-2
TOEP(4)	SUBID	0.01872	0.007469	2.51	0.0122	Residual covariance for t-3
Residual		0.1702	0.01401	12.15	<.0001	Residual e variance (equal across weeks)

Fit Statistics

-2 Res Log Likelihood	927.3
AIC (smaller is better)	937.3
AICC (smaller is better)	937.4
BIC (smaller is better)	951.4

Predicted Total V Matrix from Random Intercept (G) + 3-Lag TOEP Covariance (R):

Estimated V Matrix for SUBID 100 → TOTAL variance and covariance put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.8097**	0.7252	0.6974	0.6582	0.6395	0.6395	0.6395
2	0.7252	**0.8097**	0.7252	0.6974	0.6582	0.6395	0.6395
3	0.6974	0.7252	**0.8097**	0.7252	0.6974	0.6582	0.6395
4	0.6582	0.6974	0.7252	**0.8097**	0.7252	0.6974	0.6582
5	0.6395	0.6582	0.6974	0.7252	**0.8097**	0.7252	0.6974
6	0.6395	0.6395	0.6582	0.6974	0.7252	**0.8097**	0.7252
7	0.6395	0.6395	0.6395	0.6582	0.6974	0.7252	**0.8097**

Original observed (Unstructured R) Matrix (i.e., what it's trying to match with fewer parameters):

Estimated R Matrix for SUBID 100 → TOTAL variance & covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7477**	0.7190	0.6817	0.6807	0.6124	0.6182	0.6328
2	0.7190	**0.8266**	0.7428	0.7475	0.6571	0.6619	0.6541
3	0.6817	0.7428	**0.8346**	0.7498	0.6524	0.6390	0.6172
4	0.6807	0.7475	0.7498	**0.8366**	0.7292	0.7206	0.6805
5	0.6124	0.6571	0.6524	0.7292	**0.8285**	0.7280	0.7313
6	0.6182	0.6619	0.6390	0.7206	0.7280	**0.7911**	0.7352
7	0.6328	0.6541	0.6172	0.6805	0.7313	0.7352	**0.8464**

One last thing to check – do the residual variances in **R** need to be heterogeneous?

```
TITLE 'SAS Random Intercept + Lag-3H TOEP Correlation Model';
PROC MIXED DATA=&datafile. COVTEST NOCLPRINT NOITPRINT METHOD=REML;
      CLASS subid week;
      MODEL severity = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / V VCORR G TYPE=UN SUBJECT=subid;
      REPEATED week / TYPE=TOEPH(4) R RCORR SUBJECT=subid; RUN;
```

Estimated R Matrix for SUBID 100 → WP residual variance & covariance

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.1261**	0.06946	0.06022	0.01485			
2	0.06946	**0.1462**	0.09301	0.05131	0.01985		
3	0.06022	0.09301	**0.2261**	0.09152	0.07918	0.02044	
4	0.01485	0.05131	0.09152	**0.1416**	0.08986	0.05187	0.01808
5		0.01985	0.07918	0.08986	**0.2180**	0.09231	0.07197
6			0.02044	0.05187	0.09231	**0.1494**	0.08546
7				0.01808	0.07197	0.08546	**0.1868**

Note that the *covariances are unequal* within a band (because the variances are allowed to differ over time), but the *correlations are equal* within a band (in the RCORR matrix, below).

Estimated R Correlation Matrix for SUBID 100 → WP residual correlation

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	1.0000	0.5115	0.3566	0.1112			
2	0.5115	1.0000	0.5115	0.3566	0.1112		
3	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112	
4	0.1112	0.3566	0.5115	1.0000	0.5115	0.3566	0.1112
5		0.1112	0.3566	0.5115	1.0000	0.5115	0.3566
6			0.1112	0.3566	0.5115	1.0000	0.5115
7				0.1112	0.3566	0.5115	1.0000

Estimated G Matrix → BP random intercept variance

Row	Effect	subject id	Col1
1	Intercept	100	0.6503

V Matrix from Random Intercept (**G**) + 3-Lag TOEP Correlation with Heterogeneous Variances (**R**):

Estimated V Matrix for SUBID 100 → TOTAL variance and covariance put back together

Row	Col1	Col2	Col3	Col4	Col5	Col6	Col7
1	**0.7764**	0.7198	0.7105	0.6652	0.6503	0.6503	0.6503
2	0.7198	**0.7965**	0.7433	0.7016	0.6701	0.6503	0.6503
3	0.7105	0.7433	**0.8764**	0.7418	0.7295	0.6707	0.6503
4	0.6652	0.7016	0.7418	**0.7919**	0.7402	0.7022	0.6684
5	0.6503	0.6701	0.7295	0.7402	**0.8683**	0.7426	0.7223
6	0.6503	0.6503	0.6707	0.7022	0.7426	**0.7997**	0.7358
7	0.6503	0.6503	0.6503	0.6684	0.7223	0.7358	**0.8371**

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	SUBID	0.6503	0.09069	7.17	<.0001	Random intercept U variance
Var(1)	SUBID	0.1261	0.02255	5.59	<.0001	Residual e Variance at week 1
Var(2)	SUBID	0.1462	0.02283	6.41	<.0001	Residual e Variance at week 2
Var(3)	SUBID	0.2261	0.03247	6.96	<.0001	Residual e Variance at week 3
Var(4)	SUBID	0.1416	0.02442	5.80	<.0001	Residual e Variance at week 4
Var(5)	SUBID	0.2180	0.03228	6.75	<.0001	Residual e Variance at week 5
Var(6)	SUBID	0.1494	0.02411	6.20	<.0001	Residual e Variance at week 6
Var(7)	SUBID	0.1868	0.02890	6.47	<.0001	Residual e Variance at week 7
TOEPH(1)	SUBID	0.5115	0.03899	13.12	<.0001	Residual correlation for t-1
TOEPH(2)	SUBID	0.3566	0.03764	9.48	<.0001	Residual correlation for t-2
TOEPH(3)	SUBID	0.1112	0.04115	2.70	0.0069	Residual correlation for t-3

Fit Statistics

-2 Res Log Likelihood	910.0
AIC (smaller is better)	932.0
AICC (smaller is better)	932.3
BIC (smaller is better)	963.0

The RI and homogeneous variance TOEP(4) 3-lag model is nested within the RI and heterogeneous TOEP(4) 3-lag model, so we can compare deviances: $-2\Delta LL(6) = 17.3$, $p = .008$. Yes, separate variances improves model fit (AIC agrees, but BIC does not). One could go either way… in that case I'd probably re-check after adding predictors to see if it makes a difference – if not, it's parsimony vs. fit – your call.